

Hg

ENTERED



PCT

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/762,568A

DATE: 03/27/2003

TIME: 13:23:06

Input Set : N:\CrF4\03272003\I762568.raw  
 Output Set: N:\CRF4\03272003\I762568A.raw

```

1 <110> APPLICANT: Nippon Institute for Biological Science
2 <120> TITLE OF INVENTION: novel plasmid vector
3 <130> FILE REFERENCE: PCTF0001-0
4 <140> CURRENT APPLICATION NUMBER: US/09/762,568A
5 <141> CURRENT FILING DATE: 2001-02-06
6 <150> PRIOR APPLICATION NUMBER: JP, Japanese Patent
7 <151> PRIOR FILING DATE: 1999-06-04
8 <160> NUMBER OF SEQ ID NOS: 13
9 <170> SOFTWARE: FastSEQ for Windows Version 4.0
11 <210> SEQ ID NO: 1
12 <211> LENGTH: 31
13 <212> TYPE: DNA
14 <213> ORGANISM: Artificial Sequence
15 <220> FEATURE:
16 <223> OTHER INFORMATION: Designed PCR primer including 3' region of U3 and
17 VspI restriction enzyme site to multiply RSV LTR.
18 <400> SEQUENCE: 1
19 ggcattaaatg tagtcttatg caataactcct g 31
21 <210> SEQ ID NO: 2
22 <211> LENGTH: 40
23 <212> TYPE: DNA
24 <213> ORGANISM: Artificial Sequence
25 <220> FEATURE:
26 <223> OTHER INFORMATION: Designed PCR primer including 5' non coding region
27 of p19 gene, HincII, EcoRV and BglII restriction
28 enzyme site to multiply RSV LTR and down stream
29 region of LTR.
30 <400> SEQUENCE: 2
31 gttaacgata tcagatctgc ttgatccacc gggcgaccag 40
33 <210> SEQ ID NO: 3
34 <211> LENGTH: 36
35 <212> TYPE: DNA
36 <213> ORGANISM: Artificial Sequence
37 <220> FEATURE:
38 <223> OTHER INFORMATION: Designed PCR primer including 5' region of RSV
39 integrase gene and BamHI restriction enzyme site
40 to multiply RSV integrase gene.
41 <400> SEQUENCE: 3
42 ttggatccat gcccttgaga gaggttaaag atcttc 36
44 <210> SEQ ID NO: 4
45 <211> LENGTH: 33
46 <212> TYPE: DNA
47 <213> ORGANISM: Artificial Sequence

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/762,568A

DATE: 03/27/2003

TIME: 13:23:06

Input Set : N:\Crf4\03272003\I762568.raw  
 Output Set: N:\CRF4\03272003\I762568A.raw

```

48 <220> FEATURE:
49 <223> OTHER INFORMATION: Designed PCR primer including 3' region of RSV
50     integrase gene, polyA signal to multiply RSV
51     integrase gene.
52 <400> SEQUENCE: 4
53     tttatttaa ctctcggtgg cagcaagggt gtc          33
55 <210> SEQ ID NO: 5
56 <211> LENGTH: 29
57 <212> TYPE: DNA
58 <213> ORGANISM: Artificial Sequence
59 <220> FEATURE:
60 <223> OTHER INFORMATION: Designed PCR primer including 5' region of U5 and
61     VspI restriction enzyme site to multiply RSV LTR.
62 <400> SEQUENCE: 5
63     ggcattaatg aagccttctg cttcattca          29
65 <210> SEQ ID NO: 6
66 <211> LENGTH: 51
67 <212> TYPE: DNA
68 <213> ORGANISM: Artificial Sequence
69 <220> FEATURE:
70 <223> OTHER INFORMATION: Designed PCR primer including 3' region of RSV
71     integrase gene, polyA signal, nuclear localization
72     signal of SV40 large T antigen to multiply RSV
73     integrase gene.
74 <400> SEQUENCE: 6
75     tttatttaa accttcctct tcttcttagg actctcggttg gcagcaagggt t      51
77 <210> SEQ ID NO: 7
78 <211> LENGTH: 858
79 <212> TYPE: DNA
80 <213> ORGANISM: Rous sarcoma virus
81 <220> FEATURE:
82 <221> NAME/KEY: TATA_signal
83 <222> LOCATION: (84)...(90)
84 <220> FEATURE:
85 <221> NAME/KEY: polyA_signal
86 <222> LOCATION: (107)...(112)
87 <220> FEATURE:
88 <221> NAME/KEY: TATA_signal
89 <222> LOCATION: (431)...(437)
90 <220> FEATURE:
91 <221> NAME/KEY: polyA_signal
92 <222> LOCATION: (454)...(459)
93 <223> OTHER INFORMATION: A part of circular form of RSV DNA, tandem repeat
94     LTRs and adjacent non coding region.
95 <400> SEQUENCE: 7
96     acgatcggtgc cttatttagga agccaacaga cgggtctaac acggattgga cgaaccactg 60
97     aattccgcat tgccggatata ttgtatttaa gtgcctagct cgataacaata aacgccattt 120
98     taccattcac cacattgttg tgcacctggg ttgatggctg gaccgttgat tccctgacga 180
99     ctacgagcac atgcatgaag cagaaggctt cattaatgta gtcttatgca atactcctgt 240

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/762,568A

DATE: 03/27/2003

TIME: 13:23:06

Input Set : N:\Crf4\03272003\I762568.raw  
 Output Set: N:\CRF4\03272003\I762568A.raw

```

100 agtcttgcaa catgcttatg taacgatgag ttagcaacat gccttacaag gagagaaaag 300
101 gcaccgtgca cgacgattgg tgaagtaag gtggtatgat cgttagtacg atcgtgcctt 360
102 attaggaagg caacagacgg gtctaacacg gattggacga accactgaat tccgcattgc 420
103 ggagatattg tatttaagtg cctagctcgta tacaataaac gccatttac cattcaccac 480
104 atttgtgtgc acctgggtt atggctggac cgttgattcc ctgacgacta cgagcacatg 540
105 catgaagcag aaggcttcat ttggtgaccc cgacgtgatc gttaggaaat agtggtcggc 600
106 cacagacggc gtggcgatcc tgccctcattc cgtctcgctt attcggggag cggacgatga 660
107 ccctagtaga gggggctgcg gcttaggagg gcagaagctg agtggcgtcg gagggagctc 720
108 tactgcaggg agccccagata ccctaccgag aactcagaga gtcgttgaa gacgggaaga 780
109 aagcccacg actgagcggt ccaccccgagg cgtgattccg gttgctctgc gtgaccctgg 840
110 tcgcccgggtg gatcaagc 858
112 <210> SEQ ID NO: 8
113 <211> LENGTH: 972
114 <212> TYPE: DNA
115 <213> ORGANISM: Rous sarcoma virus
116 <220> FEATURE:
117 <221> NAME/KEY: CDS
118 <222> LOCATION: (1)...(972)
119 <223> OTHER INFORMATION: precursor integrase or p36 protein
120 <220> FEATURE:
121 <221> NAME/KEY: CDS
122 <222> LOCATION: (1)...(858)
123 <223> OTHER INFORMATION: mature integrase or p32 protein
124 <400> SEQUENCE: 8
125 ccc ttg aga gag gct aaa gat ctt cat acc gct ctc cat att gga ccc 48
126 Pro Leu Arg Glu Ala Lys Asp Leu His Thr Ala Leu His Ile Gly Pro
127 1 5 10 15
128 cgc gcg cta tcc aaa gcg tgt aat ata tct atg cag cag gct agg gag 96
129 Arg Ala Leu Ser Lys Ala Cys Asn Ile Ser Met Gln Gln Ala Arg Glu
130 20 25 30
131 gtt gtt cag acc tgc ccg cat tgt aat tca gcc cct gcg ttg gag gcc 144
132 Val Val Gln Thr Cys Pro His Cys Asn Ser Ala Pro Ala Leu Glu Ala
133 35 40 45
134 gga gta aac cct agg ggt ttg gga ccc cta cag ata tgg cag aca gac 192
135 Gly Val Asn Pro Arg Gly Leu Gly Pro Leu Gln Ile Trp Gln Thr Asp
136 50 55 60
137 ttt acg ctt gag cct aga atg gcc ccc cgt tcc tgg ctc gct gtt act 240
138 Phe Thr Leu Glu Pro Arg Met Ala Pro Arg Ser Trp Leu Ala Val Thr
139 65 70 75 80
140 gtg gac acc gcc tca tca gcg ata gtc gta act cag cat ggc cgt gtc 288
141 Val Asp Thr Ala Ser Ser Ala Ile Val Val Thr Gln His Gly Arg Val
142 85 90 95
143 aca tcg gtt gta caa cat cat tgg gcc acg gct atc gcc gtt ttg 336
144 Thr Ser Val Ala Val Gln His His Trp Ala Thr Ala Ile Ala Val Leu
145 100 105 110
146 gga aga cca aag gcc ata aaa aca gat aac ggg tcc tgc ttc acg tct 384
147 Gly Arg Pro Lys Ala Ile Lys Thr Asp Asn Gly Ser Cys Phe Thr Ser
148 115 120 125
149 aaa tcc acg cga gag tgg ctc gcg aga tgg ggg ata gca cac acc acc 432

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/762,568A

DATE: 03/27/2003

TIME: 13:23:06

Input Set : N:\Crf4\03272003\I762568.raw  
 Output Set: N:\CRF4\03272003\I762568A.raw

150	Lys Ser Thr Arg Glu Trp Leu Ala Arg Trp Gly Ile Ala His Thr Thr	
151	130 135 140	
152	ggg att ccg ggt aat tcc cag ggt caa gct atg gta gag cggtt ggc aac	480
153	Gly Ile Pro Gly Asn Ser Gln Gly Gln Ala Met Val Glu Arg Ala Asn	
154	145 150 155 160	
155	cgg ctc ctg aaa gat agg atc cgt gtg ctt gcg gag ggg gac ggc ttt	528
156	Arg Leu Leu Lys Asp Arg Ile Arg Val Leu Ala Glu Gly Asp Gly Phe	
157	165 170 175	
158	atg aaa aga atc ccc acc agc aaa cag ggg gaa cta tta gcc aag gca	576
159	Met Lys Arg Ile Pro Thr Ser Lys Gln Gly Glu Leu Leu Ala Lys Ala	
160	180 185 190	
161	atg tat gcc ctc aat cac ttt gag cgt ggt gaa aac acg aaa aca ccg	624
162	Met Tyr Ala Leu Asn His Phe Glu Arg Gly Glu Asn Thr Lys Thr Pro	
163	195 200 205	
164	ata caa aaa cac tgg aga cct acc gtt ctt aca gaa gga ccc ccg gtt	672
165	Ile Gln Lys His Trp Arg Pro Thr Val Leu Thr Glu Gly Pro Pro Val	
166	210 215 220	
167	aaa ata cga ata gag aca ggg gag tgg gaa aaa gga tgg aac gtg ctg	720
168	Lys Ile Arg Ile Glu Thr Gly Glu Trp Glu Lys Gly Trp Asn Val Leu	
169	225 230 235 240	
170	gtc tgg gga cga ggt tat gcc gct gtg aaa aac agg gac act gat aag	768
171	Val Trp Gly Arg Gly Tyr Ala Ala Val Lys Asn Arg Asp Thr Asp Lys	
172	245 250 255	
173	gtt att tgg gta ccc tct cga aaa gtt aaa ccg gac atc acc caa aag	816
174	Val Ile Trp Val Pro Ser Arg Lys Val Lys Pro Asp Ile Thr Gln Lys	
175	260 265 270	
176	gat gag gtg act aag aaa gat gag gcg agc cct ctt ttt gca ggc att	864
177	Asp Glu Val Thr Lys Lys Asp Glu Ala Ser Pro Leu Phe Ala Gly Ile	
178	275 280 285	
179	tct gac tgg ata ccc tgg gga gac aag caa gaa gga ctc caa gga gaa	912
180	Ser Asp Trp Ile Pro Trp Gly Asp Lys Gln Glu Gly Leu Gln Gly Glu	
181	290 295 300	
182	acc gct agc aac aag caa gaa aga ccc gga gaa gac acc ctt gct gcc	960
183	Thr Ala Ser Asn Lys Gln Glu Arg Pro Gly Glu Asp Thr Leu Ala Ala	
184	305 310 315 320	
185	aac gag agt taa	972
186	Asn Glu Ser *	
188	<210> SEQ ID NO: 9	
189	<211> LENGTH: 21	
190	<212> TYPE: DNA	
191	<213> ORGANISM: Artificial Sequence	
192	<220> FEATURE:	
193	<223> OTHER INFORMATION: Designed PCR primer including 5' region of GFP gene and a part of NheI restriction enzyme site to multiply GFP gene.	
196	<400> SEQUENCE: 9	
197	ctagcgctac cgggtcgccac c	21
199	<210> SEQ ID NO: 10	
200	<211> LENGTH: 20	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/762,568A

DATE: 03/27/2003

TIME: 13:23:06

Input Set : N:\Crf4\03272003\I762568.raw  
Output Set: N:\CRF4\03272003\I762568A.raw

201 <212> TYPE: DNA  
202 <213> ORGANISM: Artificial Sequence  
203 <220> FEATURE:  
204 <223> OTHER INFORMATION: Designed PCR primer including antisense sequence  
205 of GFP ORF to multiply a part of GFP gene.  
206 <400> SEQUENCE: 10  
207 gttgccgtcc tccttgaagt 20  
209 <210> SEQ ID NO: 11  
210 <211> LENGTH: 21  
211 <212> TYPE: DNA  
212 <213> ORGANISM: Artificial Sequence  
213 <220> FEATURE:  
214 <223> OTHER INFORMATION: Designed PCR primer including U5 region LTR  
215 sequence to  
216 multiply a part of integrated plasmid vector.  
217 <400> SEQUENCE: 11  
218 ttggtgtgca cctgggttga t 21  
220 <210> SEQ ID NO: 12  
221 <211> LENGTH: 36  
222 <212> TYPE: DNA  
223 <213> ORGANISM: Artificial Sequence  
224 <220> FEATURE:  
225 <223> OTHER INFORMATION: Designed PCR primer including 5' end of GFP ORF  
226 sequence to multiply a part of GFP gene.  
227 <400> SEQUENCE: 12  
228 atggtgagca agggcgagga gctgttcacc ggggtg 36  
230 <210> SEQ ID NO: 13  
231 <211> LENGTH: 20  
232 <212> TYPE: DNA  
233 <213> ORGANISM: Artificial Sequence  
234 <220> FEATURE:  
235 <223> OTHER INFORMATION: Designed PCR primer including a part of GFP ORF  
236 sequence to multiply a part of GFP gene.  
237 <400> SEQUENCE: 13  
238 gtcgagctgg acggcgacgt 20

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/762,568A

DATE: 03/27/2003

TIME: 13:23:07

Input Set : N:\Crf4\03272003\I762568.raw

Output Set: N:\CRF4\03272003\I762568A.raw